

Covariance

Covariance is a measure of covariation between two data sets. It is a measure of how closely two variables are associated with another. Specifically, covariance is the sum of the product of the deviations. It can be displayed mathematically as:

$$s_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

s_{xy} = covariance between x and y

\bar{x} = average of variable x

\bar{y} = average of variable y

Covariance is similar to the correlation coefficient; however covariance is linear and it does not vary between -1 and +1. In fact, the correlation coefficient between two variables is equal to the covariance between those variables divided by the product of their standard deviations.

$$r = \frac{s_{xy}}{s_x s_y}$$

r = correlation coefficient between x and y

s_{xy} = covariance between x and y

s_x = standard deviation of x

s_y = standard deviation of y